



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes

### Citation for published version:

Charlesworth, B, Jordan, CY & Charlesworth, D 2014, 'The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes', *Evolution: International Journal of Organic Evolution*. <https://doi.org/10.1111/evo.12364>

### Digital Object Identifier (DOI):

[10.1111/evo.12364](https://doi.org/10.1111/evo.12364)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Evolution: International Journal of Organic Evolution

### Publisher Rights Statement:

Open Access

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



**The evolutionary dynamics of sexually antagonistic mutations in  
pseudoautosomal regions of sex chromosomes**

Brian Charlesworth, Crispin Y. Jordan and Deborah Charlesworth

Institute of Evolutionary Biology, School of Biological Sciences

The University of Edinburgh, Ashworth Lab., King's Buildings

West Mains Road, Edinburgh EH9 3JT, UK

Corresponding author: Deborah Charlesworth

email: Deborah Charlesworth <Deborah.Charlesworth@ed.ac.uk>

Keywords: Sexual antagonism, sex chromosomes, pseudo-autosomal region, recombination,  
gene expression

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/evo.12364.

Accepted Article

Sex chromosomes can evolve gene contents that differ from the rest of the genome, as well as larger sex differences in gene expression compared with autosomes. This probably occurs because fully sex-linked beneficial mutations substitute at different rates from autosomal ones, especially when fitness effects are sexually antagonistic (SA). The evolutionary properties of genes located in the recombining pseudo-autosomal region (PAR) of a sex chromosome have not previously been modelled in detail. Such PAR genes differ from classical sex-linked genes by having two alleles at a locus in both sexes; in contrast to autosomal genes, however, variants can become associated with gender. The evolutionary fates of PAR genes may therefore differ from those of either autosomal or fully sex-linked genes. Here, we model their evolutionary dynamics by deriving expressions for the selective advantages of PAR gene mutations under different conditions. We show that, unless selection is very strong, the probability of invasion of a population by an SA mutation is usually similar to that of an autosomal mutation, unless there is close linkage to the sex-determining region. Most PAR genes should thus evolve similarly to autosomal rather than sex-linked genes, unless recombination is very rare in the PAR.

Key words: Sexually antagonistic selection, adaptation, selective advantage, partial sex-linkage, pseudo-autosomal region, gene expression

Accepted Article

The evolutionary fates of mutant alleles may differ between sex chromosomes and autosomes. For example, in an XY system, a male-benefit mutant allele of a Y-linked gene can spread even if it is potentially lethal to females, whereas invasion of a population by a beneficial autosomal mutation requires a larger fitness benefit to males than any cost to females. Most modelling work on this aspect of sex chromosome evolution has focused on a fully sex-linked region with no recombination in the heterogametic sex, with a “genetically degenerate” Y chromosome that has lost nearly all genes present on the X, so that males are generally hemizygous for X-linked loci. In addition to any differences in effective population size that can affect the substitution rates of weakly selected mutations, two major factors can cause differences between the evolution of fully sex-linked regions and autosomes: opposing selection pressures on the two sexes (sexual antagonism, or “SA”), and the dominance or recessivity of the fitness effects of adaptive mutations; see reviews by Vicoso and Charlesworth (2006), Vicoso and Charlesworth (2009), Connallon and Clark (2010) and Meisel and Connallon (2013). For brevity, we will summarize previous results with reference to XY systems; ZW systems mostly behave similarly, interchanging males and females (in the Discussion, we mention situations where XY and ZW systems are expected to differ).

Under male hemizygosity, a new mutation in an X-linked gene is present two-thirds of the time in females, but only one-third of the time in males. Thus, other things being equal, a fully dominant female-benefit mutant allele at an X-linked locus can invade a population unless male fitness is reduced by twice its effect on females (Rice 1984). Favorable mutations with full or partial dominance at fully X-linked genes (with no Y-linked alleles) are therefore more likely than autosomal genes to evolve enhanced female, relative to male, functions. However, strongly recessive favorable mutations affecting males can be established more easily on the X chromosome than the autosomes, because of their greater exposure to

Accepted Article

selection in hemizygous males when rare, and their fixation could result in the opposite pattern to that just described (Rice 1984; Charlesworth et al. 1987). There has been much discussion of how well these predictions are supported by studies of genome-wide patterns of gene expression, which are expected to correlate with effects on fitness (see the reviews cited above; also Mank *et al.* 2010; Meisel, Malone and Clark 2012).

The models just outlined have not considered loci with functional homologs on both the X and Y chromosomes. This situation can arise in fully sex-linked regions when there has not been enough time for major genetic degeneration of the Y or W chromosome. X-linked genes with functional Y-linked alleles can also exist when the genetic degeneration of a Y-linked gene is prevented by selection in the haploid stage of the life cycle, e.g. in predominantly haploid plants such as liverworts (Okada et al. 2001), and in the male gametophytes in pollen in flowering plants (Bergero and Charlesworth 2011; Chibalina and Filatov 2011).

The recombining, partially sex-linked pseudoautosomal regions (“PARs”; see Figure 1) of many extant XY or ZW systems also carry functional copies on both sex chromosomes.

While recombination continues, PAR genes will not undergo genetic degeneration, and both sexes will have two alleles at PAR loci. It is therefore of interest to understand how such genes will be expected to evolve. The evolution of these regions has been little studied (Otto et al. 2011), although it is clear that non-recombining regions of sex chromosomes have repeatedly evolved from states with larger recombining pseudo-autosomal regions in Eutherian mammals (Lahn and Page 1999; Skaletsky et al. 2003), birds (Lawson-Handley et al. 2004; Nishida-Umehara et al. 2007; Nanda et al. 2008; Pigozzi 2011), snakes (Matsubara et al. 2006; Vicoso et al. 2013a) and plants (Bergero et al. 2007; Wang et al. 2012). New partially sex-linked genome regions can also arise by chromosome fusions or translocations that add autosomal regions to the X chromosome, provided that the added region continues to recombine with the homologous autosome in males (reviewed in Bachtrog 2012).

Accepted Article

The strength of selection for male- and female-benefit alleles of partially sex-linked genes has not yet been studied quantitatively. We therefore examine the fates of new mutations in a pseudoautosomal region of a sex chromosome system. We ask whether PAR genes differ from autosomal loci in their tendency to fix alleles with sex-specific fitness effects (as fully sex-linked loci do). Specifically, we relate the net strength of selection on SA mutations to the frequency of recombination with the sex-determining region. The model also yields results for fully X- and non-degenerated Y-linked alleles, which is the extreme of the situation for PAR genes, with no recombination between X and Y.

Our model is described in detail in the next section. It relates to two evolutionary situations that might result in sex differences in gene expression. A mutation at an SA locus may affect gene expression differently in the two sexes (for example, due to their hormonal state or other consequences of their different genetic backgrounds), and may spread to intermediate frequencies or fixation (Connallon and Clark 2010), leading directly to sex-specific gene expression. Alternatively, sex-specific gene expression may be “secondary” effects, evolved in response to the spread of alleles with SA effects (whether these arise from expression differences or in another way). Reduced expression of such genes in males might then “resolve” the conflict and create appropriate sex-specific gene expression (Rice 1984; Vicoso and Charlesworth 2006); the resulting removal of the harmful fitness effects of such mutations will result in their fixation, leaving a sex difference in gene expression controlled by the second change. The converse pattern of sex-specific change in expression would occur with partially recessive mutations, as discussed above. In either case, our question is: what portion of a PAR should show a greater tendency to accumulate SA alleles, compared with autosomal loci? As one might expect intuitively, we show that this effect is likely to be confined to PAR genes closely linked to the boundary with the sex-determining region.

### *The Models and Analytical Results*

We use exact and approximate analytic expressions to study the invasion of populations by either male- or female-benefit alleles. As in previous models of X-linked genes with male hemizygosity, we assume that, for the fully sex-linked region, males are always XY and females XX (i.e. the YY genotype is inviable). We study selection coefficients of 0.1 or less, to which the approximations used below should apply reasonably well. Both theory on the fixation of mutations during “adaptive walks” (Orr 2005), and evidence from studies of molecular evolution in *Drosophila* (Sella et al. 2009; Schneider et al. 2011), suggest that small fitness effects will predominate for favorable mutations.

We concentrate on the conditions for invasion of a population by a newly arisen mutant allele, given the different forces acting on it. We quantify the net selective advantage,  $\sigma$ , of a rare allele by the asymptotic value of its rate of spread into the population ( $\sigma \sim \Delta x / x$ , where  $x \ll 1$  is the allele frequency in question). This is closely related to the probability of survival of a new mutation (which is equivalent to its probability of fixation when a balanced polymorphism cannot be maintained). In the case of an autosomal or fully sex-linked mutation, the survival probability is approximately  $2\sigma$  in a Wright-Fisher population whose population size  $N$  is sufficiently large that  $Ns > 1$  (Haldane 1927; Charlesworth et al. 1987). The asymptotic rate of spread ignores the need to consider the three possible origins of mutations (X-linked in a female, X-linked in a male, and Y-linked in a male), whose survival probabilities differ and should be appropriately averaged. However, the analysis described in the ESM shows that  $2\sigma$  provides a good approximation to the largest of the three survival probabilities (see Tables S1 and S2). It also shows that, overall, our results described below overestimate the effect of reducing recombination in increasing the net survival probability of a new PAR mutation, even when allowing the possibility of a higher mutation rate in males than females (see ESM). By deriving expressions for  $\sigma$  for both male- and female-benefit alleles, we also obtain analytical approximations for the parameter values under which

balanced polymorphisms can be maintained.

## MALE-BENEFIT MUTATIONS

**The fitness scheme:** We assume an invading male-benefit allele,  $A_2$ , and a resident female-benefit allele,  $A_1$ . The recombination fraction between the A locus and the sex-determining region is denoted by  $r$  (Figure 1). Using a modification of the fitness scheme of Jordan and Charlesworth (2012), the fitnesses in the two sexes are specified by three parameters:  $s$ , the selection coefficient against the less fit homozygous genotype in males ( $A_1A_1$ ); a constant,  $c$ , that represents the relative strength of selection in females versus males; the dominance coefficient,  $h$ , with respect to the fitness effect of  $A_1$  in males and of  $A_2$  in females, such that  $0 < h \leq 1$  ( $h$  here thus has a different meaning from that in Jordan and Charlesworth 2012). To study invasion by  $A_2$ , only the ratios for each sex of the fitness of  $A_1A_2$  to that of the resident homozygote  $A_1A_1$  need be considered (see Table 1 and ESM). It is important to note the difference between the net selective advantage  $\sigma$  and  $s$ , the latter being the strength of selection acting on alleles at the A locus.

Table 1 here

The value of  $\sigma$  is obtained by subtracting one from the leading eigenvalue of the matrix that describes the recursion relations for the frequencies of the various possible rare genotypes carrying  $A_2$ , which are given in the ESM. While the full characteristic equation of this matrix is a cubic, and hence is hard to handle (Bull 1983, pp. 265-269), it can be reduced to a quadratic equation by assuming that  $\sigma^3$  can be ignored (see the Appendix). If terms higher than second-order in  $s$  are also ignored in the relevant solution, consistent with our assumption of weak selection, the following simple expression for  $\sigma$  is obtained, provided that  $r$  is not  $\ll s$ :



$$S = \frac{s}{2}[1 - h(1 + c)] + \frac{s^2}{8}[3 - 2h(1 + c - 2hc) - h^2(1 + c)^2] + \frac{s^2}{16r}[1 - h(1 - c)]^2 + o(s^2)$$

(1)

For  $r = 1/2$ , this is the same as the expression as for an autosomal locus, where the terms in  $s^2$  reduce to  $1/2s^2(1 - h)$ .

This expression shows that  $\sigma$  is only weakly dependent on  $r$  in this region of parameter space. The first two terms on the right hand side are independent of  $r$ , whereas the third term decreases with  $r$ . As expected intuitively,  $\sigma$  therefore increases as  $r$  decreases. However,  $r$  only affects a term of order  $s^2$ ; recombination is thus likely have an important effect only when selection is strong, except when  $1 - h(1 + c)$  is of order  $s$  or less, in which case  $\sigma$  is itself of order  $s^2$  in the range of  $r$  values where this approximation is valid.

A heuristic argument that also leads to this result is as follows. When  $r$  and  $s$  are of comparable magnitude, the effect of linkage on the selective advantage of an SA allele, increasing its value over that with free recombination, is given by the product of its degree of association with the sex that it benefits with its advantage in that sex. The latter is  $O(s)$ ; the former is  $O(s/r)$ , since the amount of linkage disequilibrium in a two-locus system is of the order of the ratio of the measure of epistasis in fitness (here of order  $s$ ) to the recombination rate (e.g. Charlesworth and Charlesworth 2010, p.420).

When  $|1 - h(1 + c)| \gg s$ , the second term on the right of equation (1) can be ignored, and the equation can be rewritten in terms of the scaled variables  $\tilde{S} = S/s$  and  $\tilde{r} = r/s$ , yielding the following expression:

$$\tilde{S} = \frac{1}{2}[1 - h(1 + c)] + \frac{1}{16\tilde{r}}[1 - h(1 - c)]^2 + o(s) \quad (2)$$

When this approximation is valid,  $\tilde{S}$  is a hyperbolic function of  $\tilde{r}$ , and its value depends on parameters that are independent of the strength of selection,  $s$ . Since both  $1 - h(1 + c)$  and

the magnitudes of the multiplier of  $1/\tilde{r}$  in equation (2) are of the order of 1,  $\tilde{S}$  is unlikely to greatly exceed its minimum value unless  $\tilde{r}$  is  $< 1$ , i.e.  $r$  is smaller than  $s$ .

When  $h(1 + c) - 1 \gg s$ , the right-hand side of equation (2) is negative, and so a loosely linked allele will not invade; invasion by a male-beneficial allele is then possible only if the condition for invasion at  $r = 0$  is satisfied (given by equation (4) below), and when  $\tilde{r}$  is below a critical value. Setting  $\tilde{S}$  to 0 in equation (3), the critical value is given by:

$$\tilde{r}_{cm} \gg \frac{[h(1 - c) - 1]^2}{8[h(1 + c) - 1]} \quad (3)$$

Since the magnitudes of the denominator and numerator of this equation are both of order 1, the critical recombination fraction is of order  $s$ , implying a requirement for very tight linkage if  $s$  is small. As discussed in the ESM, the condition for  $\sigma$  to exceed zero is equivalent to that for the survival probability of the mutation to be non-zero, so that this result also applies to the survival probability.

The above approach breaks down when  $r$  is much smaller than  $s$ , but the behavior of  $s$  as a function of  $r$  can then be investigated by evaluating  $\partial S / \partial r$  at  $r = 0$  and using this derivative in the Taylor series expansion of  $\sigma$  as a function of  $r$ , around its value at  $r = 0$  (see Appendix, equation A6). The latter value can be found as follows. With  $r = 0$ , male carriers of the invading allele have fitness  $1 - hs$ , while the fitness of the resident population is  $1 - s$ . If  $A_2$  is associated with the Y chromosome, the ratio of these two fitnesses determines its rate of spread. To terms of order  $s$ , the corresponding difference in fitness is  $s(1 - h)$ , which is equivalent to  $\sigma$ . When  $r$  is non-zero but small, we have:

$$\tilde{S} \gg (1 - h) - \tilde{r} \quad (4)$$

If  $r = 0$ , this obviously requires the mutation to have arisen on the Y chromosome. (The advantage to an X-linked mutation is smaller than this, because it is present two-thirds of the

time in females, whose fitness is reduced by the mutation.) As  $r$  increases away from zero,  $\sigma$  decreases linearly from this value if  $h < 1$ .

Equation (4) is valid even when the condition  $|1 - h(1 + c)| \gg s$  for the validity of equation (2) is not met. In this case,  $\sigma$  is of order  $s^2$  at  $r = 1/2$ , and increases to approximately  $s(1 - h)$  at  $r = 0$ , i.e.  $\sigma$  increases  $(1/s)$ -fold. Thus  $\sigma$  should decrease with increasing  $r$  over a wide range of  $r$  values, rather than becoming almost unchanging at a low  $r$  value, as occurs when  $|1 - h(1 + c)| \gg s$ ; this is confirmed by numerical examples (see below).

### FEMALE-BENEFIT MUTATIONS

In this case, an  $A_1$  allele invading an initially  $A_2A_2$  population increases females' fitness, and may be deleterious in males. The fitnesses of the heterozygotes and resident homozygotes of each sex are displayed in Table 2; we will consider only the case when the parameter  $c'$  is equal to  $c$ . Once again,  $h$  measures the dominance of the fitness effects of an allele in the sex whose fitness is reduced by its presence. The fitness of  $A_1A_2$  females relative to the resident  $A_2A_2$  population is now  $(1 - chs)/(1 - cs)$ , and that of  $A_1A_2$  males is  $1 - hs$ . If  $\sigma$  for this case and for the male-benefit case with the same set of parameters are both  $> 0$ , a polymorphism will be maintained.

Table 2 here

When  $r$  is not  $\ll s$ , a similar approach to that for a male-benefit mutation (see Appendix) results in the following approximate expression:

$$S = \frac{s}{2}[c(1 - h) - h] + \frac{s^2}{8}[3c^2 - 2hc(1 + c) - h^2(1 + c)^2] + \frac{s^2}{16r}[c(1 - h) + h]^2 + o(s^2) \quad (5)$$

As in the case of a male-benefit mutation,  $\sigma$  decreases with  $r$ , with the relevant term being of order  $s^2$ . The net selection strength when  $r$  is  $1/2$  is again the same as that for an autosomal

locus. The condition for invasion with loose linkage that allows  $\sigma$  to be positive and of order  $s$  is now  $c - h(1 + c) \gg s$ .

When  $|c - h(1 + c)| \gg s$ , we can use the same rescaling as before to obtain

$$\tilde{s} \gg \frac{1}{2}[c(1 - h) - h] + \frac{1}{16\tilde{r}}[c(1 - h) + h]^2 \quad (6)$$

This yields the following critical value of  $r/s$  for invasion by a female-benefit allele

$$\tilde{r}_{cf} \gg \frac{[c(1 - h) - h]^2}{8[h(1 + c) - c]} \quad (7)$$

(As for the male-benefit case, the condition for invasion with  $r = 0$  must also be satisfied; this is given by equation (8) below.)

When  $h > c/(1 + c)$  and  $h > 1/(1 + c)$ , a balanced polymorphism will be maintained only when the recombination fraction is smaller than the smaller of the two values given by equations (3) and (7). For the relatively high recombination rates required for (2), (3), (6) and (7) to be valid, low values of  $h$  (recessivity of the deleterious effects of the A locus mutant alleles) favor invasion by both types of mutation; in contrast, a high value of  $c$  (implying a stronger effect of the mutation in females than males) favors invasion by female-benefit mutations, but has the opposite effect on male-benefit mutations.

The value of  $\sigma$  when  $r$  is close to zero can be found by the same method as before, giving the following first-order approximation:

$$\tilde{s} = \frac{[2c(1 - h) - h]}{3} - \frac{\tilde{r}}{3} + o(s) \quad (8)$$

The first term of this expression gives  $\tilde{s}$  for the case of no recombination. This is equivalent to the value obtained by weighting the sex-specific effects of selection for a completely X-linked mutation by 2/3 for females and 1/3 for males, respectively, as is appropriate for weak selection on an X-linked allele in the case of a fully degenerate Y chromosome (Vicoso and Charlesworth 2006). The net strength of selection again declines linearly with increasing

recombination. Similarly to the male-benefit case, if the condition  $|c - h(1 + c)| \gg s$  is violated, there can be a  $(1/s)$ -fold increase in  $\sigma$  between  $r = 1/2$  and  $r = 0$ .

Comparing the conditions for invasion by male-and female benefit mutations with  $r$  near zero brings out an important contrast with the autosomal case, as well as with X-linkage with a degenerate Y chromosome. Equation (4) implies that a weakly selected male-benefit mutation will almost always be able to invade if the recombination fraction is  $\ll s$ , unless  $h$  is very close to 1. This means that it is very unlikely that a female-benefit mutation with a deleterious effect on males will spread to fixation if it is closely linked to the sex-determining region. Equation (8) implies that a male-benefit mutation will be unable to spread to fixation if  $h < 2c(1 - h)$ ; if  $c > 1/2$ , this is easily satisfied if  $h \leq 1/2$ . This is the main reason why polymorphisms for SA alleles are more likely to be maintained in the PAR than on autosomes (Jordan and Charlesworth 2012), particularly when they occur in loci closely linked to its boundary with the non-recombining region of the sex chromosomes.

### *Numerical Results*

We have generated numerical results for a range of cases, in order to test the adequacy of the approximations made above, and to illustrate the behavior of the system graphically. Figure 2 shows some results for the invasion of a population by male-benefit alleles. The parameters are defined in Table 1. We assumed that most mutations have small fitness effects, and therefore chose three values of  $s$  (which quantifies the selection against the initial male genotype), with the highest value being  $s = 0.1$ ; the figures show only the  $s = 0.1$  and  $0.01$  cases, since with  $s = 0.001$  the approximations agree very closely with the values calculated from the full equation. The figures also show results for different values of the parameter  $c$ , corresponding to stronger selection in males than females ( $c < 1$ ) and vice-versa ( $c > 1$ ),

respectively. For the harmful effects of the mutations in females, we assumed  $h = 0.5$ , because intermediate dominance is plausible for mutations with small fitness effects, and for mutations causing small changes in expression levels (see Discussion, and the discussion of the dominance of SA mutations in Connallon and Clark 2010). With larger  $h$ , invasion is less likely than in the cases shown, but the behavior is otherwise as described above.

The  $\sigma$  values calculated using the quadratic equation (A4) agree well with those using the exact characteristic equation for the system (equations A1 and A2), even when selection is strong ( $s = 0.1$ ; the results are not shown, since the curves for the two cases nearly coincide). With  $c = 0.5$ , Figure 2 shows values only for recombination rates below the assumed values of  $s$  (except for the  $r = \frac{1}{2}$  point), because, as can be seen from the figure,  $\sigma$  depends very weakly on  $r$  for larger  $r$  values. The linear approximation (4) for low  $r$  yields slightly lower values than the exact ones, but again agreement is close when  $r$  is small. As expected, equation (2) works well only for relatively loose linkage. With  $c = 2$ , the male-benefit allele cannot invade the population unless  $r$  is below a critical threshold value, and the examples shown in the figure therefore focus on  $r$  values below this value, which is well predicted by equation (A4) for both  $s$  values, but equation (3) yields a value somewhat below the correct one. Equation (2) does not work very well for the  $c = 2$  case, because non-zero  $s$  values require small  $r$  relative to  $s$ , which is where this expression becomes inaccurate.

Figure 2 here

These results confirm that equations (2) - (4) can be used to predict the fate of male benefit mutations. When equation (3) implies that there is no threshold  $r$  value, invasion is generally only slightly more likely to occur for a partially sex-linked mutation than for an autosomal one (where  $r = \frac{1}{2}$ ). Figure 2 includes the  $r = \frac{1}{2}$  case for the two selection coefficients modelled above for  $c = 0.5$ . Only mutations closely linked to the male-determining region, with  $r < s$ , show a strongly increased  $\sigma$  value. The  $\sigma$  value is less than

double that for an autosomal mutation unless  $r < 0.0038$  when  $s = 0.01$ ; for the  $s = 0.1$  case, this occurs when  $r < 0.036$ . For a 20% increase in  $\sigma$ , the corresponding  $r$  values for the two cases are about 0.015 and 0.068, respectively. We therefore conclude that PAR loci closely linked to the male-determining region are expected to exhibit a higher rate of accumulation of SA mutations, but that most genes in the PAR are not. The same conclusion applies to situations when there is a threshold  $r$  value; in such cases, the outcome for  $r$  values above the threshold is the same as for the autosomal case.

As noted in the previous section, when the condition  $|1 - h(1 + c)| \gg s$  for the validity of equation (2) is not met, equation (1) must be used for the male-benefit case; the equivalent condition for the female-benefit case is when  $|c - h(1 + c)| \gg s$  does not hold, in which case equation (5) must be used. In such cases, the expected dependence of  $\sigma$  on  $r$  is stronger at higher  $r$  values than in the cases shown above. For comparison with the  $c = 2$  case in Figure 2, a numerical example for the male-benefit case is shown in Figure 3, with  $s = 0.01$ ,  $h = 0.5$  and  $c = 1$  (so that the quantity above, as well the second term on the right of equation (1), is zero).  $\sigma$  is again much larger for small  $r$  than for  $r = 1/2$ , but it reaches double the autosomal value when  $r = 0.16$ , a much higher recombination rate than seen in the cases shown above (but still representing a small part of an entire PAR). From equation (1), it can be seen that the  $r$  value required to generate a given ratio of  $\sigma$  to its value at  $r = 1/2$  does not depend on  $s$  when  $|c - h(1 + c)| \ll s$ , so that an approximate doubling of  $\sigma$  should occur at  $r = 0.16$  for these values of  $h$  and  $c$ , over a wide range of  $s$  values.

This behavior arises because the derivatives of  $\sigma$  with respect to  $r$  in equations (1) and (5) are of order  $s^2$ ; when  $s$  is small, they therefore change very slowly until  $r$  is so small that these approximations break down, and the regions of validity of equations (4) or (8) are approached, and  $\sigma$  becomes of order  $s$  instead of  $s^2$ . This behavior, however, occurs only in a restrictive range of parameter values, so that these cases can be regarded as exceptional.

Figure 3 here

The quadratic approximation for the female-benefit case (substituting the expressions from equations (A8) into equation (A4), instead of equations (A2) for the male-benefit case) is also extremely accurate, even for  $s = 0.1$ . Figure 4 shows cases with  $c = 0.75$  and  $c = 2$ , because, with the assumed  $h$  and  $s$  values, a mutation with  $c = 0.5$  does not invade. Again, the linear approximation (equation (8) for the female-benefit case) gives good agreement with the exact value for  $r$  very close to zero. Equation (7) tends to overestimate the critical  $r$  value, the opposite of what is seen in Figure 2 for the male-benefit case. Again, therefore, the approximations can be used to gain insights into the biological situation of interest, and increased invasion probabilities (compared with autosomal mutations) are expected only for mutations closely linked to the male-determining region, unless the first term in equation (6) is of order  $s^2$ .

Figure 4 here

## *Discussion*

### **THE EVOLUTIONARY DYNAMICS OF MUTATIONS IN FULLY AND PARTIALLY SEX-LINKED GENES**

The main question previously studied in relation to PAR gene evolution was whether SA polymorphisms can be maintained, and the extent to which this is facilitated by a low recombination distance from the sex-determining region (Jordan and Charlesworth 2012). This is relevant to the question of whether recombination between X and Y (or Z and W) chromosomes has been suppressed because of SA polymorphisms in the PAR (Charlesworth et al. 2005). Our present study examines the conditions for the establishment of SA mutations in PAR genes, concentrating on the spread of weakly selected mutations, although we also provide analytical results that predict when polymorphism will occur, which agree well with



deterministic numerical calculations of the full system of genotype frequency equations for invasion by SA alleles. Although invasion conditions are not always sufficient to predict when polymorphism will result, they do so for most biologically plausible situations (Jordan and Charlesworth 2012).

To understand the extent to which PARs should display genetic differences from autosomal regions, the conditions for spread and/or fixation of SA mutations are of most interest, and such differences are potentially detectable in empirical studies (see below). Our study shows that the evolutionary dynamics of genes across most of the PAR will generally differ little from that of autosomal genes, unless it has a very short map length. Altered survival probabilities of SA mutations in PAR genes should not extend beyond genes closely linked to the boundary with the fully sex-linked region. Although we model XY systems, corresponding results apply to ZW systems, reversing the sexes; some distinctive properties of ZW systems are discussed below.

In addition, the qualitative difference between recessive and dominant mutations due to hemizyosity in the heterogametic sex has created difficulties for empirical tests of the predictions for fully sex-linked genes (see the introductory section). For genes with functional copies on both X and Y chromosomes, however, both male- and female-benefit SA mutations are disfavored if they have dominant deleterious fitness effects. Invasion of male- versus female-benefit SA mutations should therefore not depend as strongly on their level of dominance as in the case of fully sex-linked, male-hemizygous genes, except for mutations with fully recessive fitness effects. However, our results showing that most PAR genes should not evolve differently from autosomal loci suggests that empirical studies of PAR genes will not help to solve the problem of the effects of dominance on the outcome of SA selection.

## PREDICTED CHANGES IN EXPRESSION DURING SEX CHROMOSOME EVOLUTION

Most theory for sex-specific expression patterns on X chromosomes and autosomes (see Introduction) assumes that the fixation of male-benefit mutations will lead to higher levels of gene expression in males, relative to females (and vice-versa for female-benefit SA mutations). This “resolution” of SA effects by modifiers of expression will also promote the fixation of polymorphisms unless they are fully Y-linked. Our study of initial invasion conditions of SA mutations thus helps to predict whether PAR genes should evolve sex differences in gene expression, regardless of whether fixation or polymorphism occurs. They imply that the evolution of sex-specific patterns of gene expression in the PAR is not likely to differ greatly from that for autosomal genes, except for loci that are very closely linked to the sex-determining region or subject to strong selection.

## DIFFERENCES BETWEEN ZW AND XY SYSTEMS

In many species, sexual selection is stronger in males than females, potentially generating larger fitness differences among males than females, and thus leading to male-benefit mutations becoming established more easily than female-benefit mutations, other things being equal. In addition, invasion conditions are more stringent for female- than male-benefit SA mutations in the XY case, and vice versa for ZW (unless selection is much stronger on males than females), as pointed out previously (Jordan and Charlesworth 2012). This is because male-benefit mutations become strongly associated with the Y-linked male-determining region when linkage is tight, allowing them to be commoner in males than females (with  $r = 0$ , equation (4) is identical in form to that for a Y- (or W-) linked mutation with no corresponding X or Z homolog), whereas, even if a female-benefit allele is highly associated with the X chromosome, it will often be present in males (equation 8). Male-

benefit PAR mutations with small  $r$  are thus more likely to become fixed in XY systems, and female-benefit genes in ZW systems, implying that the evolution of male- rather than female-biased PAR expression for such genes is more likely to occur in XY systems, and vice-versa for ZW systems, although strong sexual selection on males could overcome this effect for ZW systems.

## EMPIRICAL RESULTS FOR PAR GENES

It is very difficult to find out whether the sex chromosomes present in a given lineage, including their PARs, have fixed unexpectedly many male- or female-benefit mutations compared with autosomal genes, and to determine whether they have SA effects. One approach has been to study sex differences in patterns of gene expression. It seems reasonable to assume that such expression differences, excluding the effects of dosage compensation, imply past sex differences in effects on fitness, or extant unresolved ones (Vicoso and Charlesworth 2006; Mank et al. 2008; Mank et al. 2010; Meisel et al. 2012), because the fixation of a male-benefit allele that harms females favors reduced expression in females, leading to the evolution of higher levels of expression in males relative to females (Rice 1984; Vicoso and Charlesworth 2006), and vice-versa for female-benefit SA alleles. This is thought to contribute to the so-called “masculinization” of X chromosomes, along with the accumulation of genes with male reproductive functions (e.g. Vicoso and Charlesworth 2006; Meisel et al. 2012; Mueller et al. 2013).

Complete Y-linkage restricts male-benefit alleles to males, partially resolving conflicts between the effects of PAR gene polymorphisms in males versus females, although the expression of X-linked female-benefit alleles can still harm males. Alternatively, rapid evolution of sex-specific expression could resolve conflicts, removing selection to reduce recombination in a PAR. It has recently been suggested that the evolution of sex-specific

expression could explain the persistence of ancient PARs, as in found in sexually monomorphic ratite birds such as emus (Vicoso et al. 2013b).

However, it is often suggested that weak sexual selection reduces the chance of selective differences between the two sexes, and hence of SA effects. Non sexually dimorphic birds, such as the polyandrous emu (Coddington and Cockburn 1995), should thus be less likely than other birds to evolve sex-specific expression of PAR genes. Comparative studies might be worthwhile, to test for associations between sexual selection or dimorphism, and sex-specific expression of PAR genes, but will require information from many species about their sex chromosomes and levels of sexual selection. In the threespine stickleback, *Gasterosteus aculeatus*, which is highly sexually dimorphic, no evidence was found for sex differences in expression of 80 PAR genes studied (Leder et al. 2010). We are not aware of any other comparable data, and PAR gene expression should be studied in other species. However, important aspects of sexual selection may not remain the same over the timescale of evolution of gene expression. Changes have occurred in the ratites, as rheas (Bruning 1973) and ostriches (Bertram 1992), are polygynous.

Assuming, however, that SA effects are important in emus, can their gene expression patterns be explained by the evolutionary dynamics of PAR mutations? An estimated 20-50% of genes in the emu PAR show higher male expression bias than autosomal genes (Vicoso et al. 2013b). This suggests the fixation of male-benefit alleles with small effects on fitness at many PAR genes, followed by the evolution of lower expression in females in response to SA effects (or simply fixations of alleles with higher expression in males, associated with enhanced male fitness). Switching the sexes, so that female-benefit alleles in the XY case correspond to male-benefit alleles in the emu ZW case, our results imply that at least 20% of the genes studied must be closely linked to the fully W-linked female-specific region. Unfortunately, there is currently no genetic map for the emu, nor any information about the

Accepted Article

locations of genes in the PAR in relation to the sites of crossing over in meiosis. However, if the physically large emu PAR indeed occupies a small genetic map length, this would suggest that this species has evolved partial recombination suppression.

Overall, it thus remains unclear why some lineages evolve sex-specific expression of SA alleles in partially sex-linked genes, while others reduce recombination with the fully sex-linked region. Possibly, some lineages simply have a low genetic variance for recombination rates, and cannot respond to selection favoring suppressed recombination.

## CONCLUSIONS AND FUTURE POSSIBILITIES

A major reason for studying the models described here is to develop predictions that may generate tests for the importance of sexually antagonistic fitness effects of mutations. Ideally, direct evidence is needed for trade-offs between the effects of mutations on different organismal functions, and detailed genetic information, such as the proportion of loci involved and the estimated strengths of selection. Although there is currently considerable evidence for SA effects in species with long-established separate sexes and sex chromosomes (e.g. Connallon and Knowles 2005; Mank and Ellegren 2009), this is indirect. No conclusive examples of such genes are known, other than some cases of sexual selection in fish species (Fisher 1930; Lindholm and Breden 2002; Gordon et al. 2012). Evidence for the differential accumulation of genes with sex-biased expression might be a further source of evidence, using genome regions where this is predicted. Alternatively, tests for balanced polymorphisms due to SA selection could use recently evolved sex chromosomes with extensive PARs. In the flowering plant *Silene latifolia*, a study of sequence polymorphisms suggested high diversity in the PAR relative to other genomic regions, consistent with such a situation (Qiu et al. 2013). Such studies have not yet been related to patterns of evolution of DNA sequences, or of gene expression.

## ACKNOWLEDGEMENTS

We thank Mark Kirkpatrick and an anonymous reviewer for their comments on the previous version of this paper. C.Y. Jordan was supported by BBSRC grant RA2044 to R.A. Ennos.

## DATA ARCHIVING

doi: 10.5061/dryad.t1m2q

## LITERATURE CITED

- Bachtrog, D., 2012. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nature Rev. Genet.* 14: 113 -124.
- Bergero, R., and D. Charlesworth, 2011. Preservation of the Y transcriptome in a 10MY old plant sex chromosome system. *Curr. Biol.* 21: 1470-1474.
- Bergero, R., A. Forrest, E. Kamau and D. Charlesworth, 2007. Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* 175: 1945-1954.
- Bertram, B., 1992. *The Ostrich Communal Nesting System*. Princeton University Press, Princeton, NJ.
- Bruning, D., 1973. The greater rhea chick and egg delivery route. *Nat. Hist.* 82: 62-75.
- Bull, J. J., 1983. *Evolution of Sex Determining Mechanisms*. Benjamin/Cummings, Menlo Park, CA.
- Charlesworth, B., J. A. Coyne and N. H. Barton, 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am.Nat.* 130: 113-146.
- Charlesworth, D., and B. Charlesworth, 1975. Theoretical genetics of Batesian mimicry. II. Evolution of supergenes. *J. Theoret. Biol.* 55: 305-324.
- Charlesworth, D., and B. Charlesworth, 2010. *Elements of Evolutionary Genetics*. Roberts and Company, Greenwood Village, Colorado.
- Charlesworth, D., B. Charlesworth and G. Marais, 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95: 118-128.
- Chibalina, M., and D. Filatov, 2011. Plant Y chromosome degeneration is retarded by haploid purifying selection. *Curr. Biol.* 21: 1475–1479.
- Coddington, C., and A. Cockburn, 1995. The mating system of free-living emus. *Australian Journal of Zoology* 43: 365–372.

- Connallon, T., and L. L. Knowles, 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends Genet.* 21: 495-499.
- Fisher, R. A., 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Gordon, S. P., A. López-Sepulcre and D. N. Reznick, 2012. Predation-associated differences in sex-linkage of wild guppy coloration. *Evol.* 66: 912-918.
- Haldane, J. B. S., 1927. A mathematical theory of natural and artificial selection. V. Selection and mutation. *Proc. Camb. Phil. Soc.* 23: 838-844.
- Jordan, C., and D. Charlesworth, 2012. The potential for sexually antagonistic polymorphism in different genome regions. *Evol.* 66: 505–516.
- Lahn, B. T., and D. C. Page, 1999. Four evolutionary strata on the human X chromosome. *Science* 286: 964-967.
- Lawson-Handley, L. J., H. Ceplitis and H. Ellegren, 2004. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics* 167: 367-376.
- Leder, E. H., J. M. Cano, T. Leinonen, P. B. O'hara, M. Nikinmaa *et al.*, 2010. Female-biased expression on the X chromosome as a key step in sex chromosome evolution in threespine sticklebacks. *Mol. Biol. Evol.* 27: 1495 - 1503.
- Lindholm, A., and F. Breden, 2002. Sex chromosomes and sexual selection in Poeciliid fishes. *Am. Nat.* 160: S214-S224.
- Mank, J. E., and H. Ellegren, 2009. Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. *Evol.* 63: 1464-1472.
- Mank, J. E., L. Hultin-Rosenberg, M. Zwahlen and H. Ellegren, 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *Am.Nat.* 171: 35–43.
- Mank, J. E., K. Nam, B. Brunström and H. Ellegren, 2010. Ontogenetic complexity of sexual dimorphism and sex-specific selection. *Mol. Biol. Evol.* 27: 1570-1578.



- Matsubara, K., H. Tarui, M. Toriba, K. Yamada, C. Nishida-Umehara *et al.*, 2006. Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. *Proc Nat Acad Sci USA* 103: 18190-18195.
- Meisel, R., J. Malone and A. Clark, 2012. Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome Res.* 22: 1255-1265.
- Meisel, R. P., and T. Connallon, 2013. The faster-X effect: integrating theory and data. *Trends Genet.* 29: 537-544.
- Mueller, J., H. Skaletsky, L. Brown, S. Zaghlul, S. Rock *et al.*, 2013. Independent specialization of the human and mouse X chromosomes for the male germ line. *Nat. Genet.* 45: 1083 - 1087.
- Nanda, I., K. Schlegelmilch, T. Haaf, M. Scharl and M. Schmid, 2008. Synteny conservation of the Z chromosome in 14 avian species (11 families) supports a role for Z dosage in avian sex determination. *Cytogenet. Genome Res.* 122: 150-156.
- Nishida-Umehara, C., Y. Tsuda, J. Ishijima, J. Ando, A. Fujiwara *et al.*, 2007. The molecular basis of chromosome orthologies and sex chromosomal differentiation in palaeognathous birds. *Chromosome Res.* 15: 721-734.
- Okada, S., T. Sone, M. Fujisawa, S. Nakayama, M. Takenaka *et al.*, 2001. The Y chromosome in the liverwort *Marchantia polymorpha* has accumulated unique repeat sequences harboring a male-specific gene. *Proc. Natl. Acad. Sci. U. S. A.* 98: 9454-9459.
- Orr, H. A., 2005. The genetic theory of adaptation: A brief history. *Nature Rev. Genet.* 6: 119-127.
- Otto, S. P., J. R. Pannell, C. L. Peichel, T. Ashman, D. Charlesworth *et al.*, 2011. About PAR: The distinct evolutionary dynamics of the pseudoautosomal region. *Trends*

Genet. 27: 358-367.

Pigozzi, M., 2011. Diverse stages of sex-chromosome differentiation in tinamid birds: evidence from crossover analysis in *Eudromia elegans* and *Crypturellus tataupa*.

Genetica 139: 771-777.

Qiu, S., R. Bergero and D. Charlesworth, 2013. Testing for the footprint of sexually antagonistic polymorphisms in the pseudo-autosomal region of a plant sex chromosome pair. Genetics 194: 663-672.

Rice, W. R., 1984. Sex chromosomes and the evolution of sexual dimorphism. Evol. 38: 735-742.

Schneider, A., B. Charlesworth, A. Eyre-Walker and P. D. Keightley, 2011. A method for inferring the rate of occurrence and fitness effects of advantageous mutations Genetics 189: 1427-1437.

Sella, G., D. Petrov, M. Przeworski and P. Andolfatto, 2009. Pervasive natural selection in the *Drosophila* genome? PLoS Genetics 5: 6.

Skaletsky, H., T. Kuroda-Kawaguchi, P. J. Minx, H. S. Cordum, L. Hillier *et al.*, 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423: 825 - 837.

Vicoso, B., and B. Charlesworth, 2006. Evolution on the X chromosome: unusual patterns and processes. Nature Rev. Genet. 7: 645-653.

Vicoso, B., and B. Charlesworth, 2009. The deficit of male-biased genes on the *D. melanogaster* X chromosome is expression-dependent: a consequence of dosage compensation? Journal of Molecular Evolution 68: 576-583.

Vicoso, B., J. Emerson, Y. Zektser, S. Mahajan and D. Bachtrog, 2013a. Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. PLoS Biol. 11: e1001643.

Vicoso, B., V. Kaiser and D. Bachtrog, 2013b. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. *Proc. Natl. Acad. Sci. U. S. A.* 110: 6453-6458.

Wang, J., J. Na, Q. Yu, A. R. Gschwend, J. Han *et al.*, 2012. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proc. Natl. Acad. Sci. U. S. A.* 109: 13710-13715.

**Table 1. The fitness model for an invading male-beneficial  $A_2$  allele in an XX/XY system**

Genotype	$A_1A_2$	$A_1A_1$
Males	$1 - hs$	$1 - s$
Females	$1 - chs$	1

Under this scheme,  $0 < c < 1$  means that selection is stronger in males than females,  $c = 1$  means that it acts equally strongly in both sexes, and  $c > 1$  implies a greater "intensity" of selection in females than males. For selection occurring only in males,  $c = 0$ ;  $c = \infty$  corresponds to selection only in females (a ZW system can be modeled by exchanging males for females). If  $c < 0$ , selection acts in the same direction in both sexes, which is of no interest for the problem under considered here.

**Table 2. The fitness model for an invading female-beneficial  $A_1$  allele in an XX/XY system**

Genotype	$A_1A_2$	$A_2A_2$
Males	$1 - hs$	1
Females	$1 - c'hs$	$1 - c's$

As before,  $c'$  measures the relative strength of selection on males and females. In general, we may have  $c' \neq c$ , since only the ratios for each sex of the fitness of the heterozygote to the fitness of the resident homozygote enter into the recursion relations (see ESM). For simplicity, however, the equations presented in the text assume  $c' = c$ , which is equivalent to assuming similar levels of dominance in each sex. They can be made completely general by inserting  $c'$  instead of  $c$ .

Figure 1. The genetic model. The figure shows an XY sex chromosome pair with a gene (indicated by a short vertical grey line) in the PAR, at a genetic map distance of  $r$  from the male-determining, or male-specific region.

Figure 1

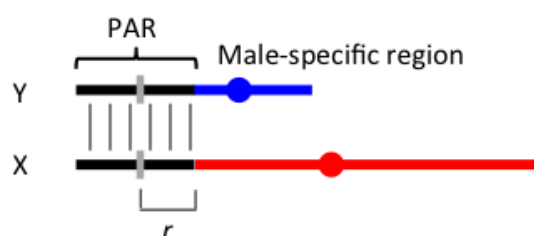


Figure 2. Results for the invasion of a population by male-benefit alleles. The y axis shows the net strength of selection ( $\sigma$ , see text), and the x axis is the frequency of recombination with the fully sex-linked region; and the  $r = 1/2$  case represents autosomal loci. The parameters are defined in Table 1. The figure shows examples of the  $s = 0.1$  and  $0.01$  cases, with  $h = 0.5$ . Results are shown for two values of the parameter  $c$  ( $0.5$  and  $2$ ), where  $c < 1$  corresponds to stronger selection in males than females and  $c > 1$  to stronger selection in females. In the former situation, the mutant allele can invade with any  $r$  value; in the latter situation, invasion occurs only if  $r$  is below a threshold value (see text).

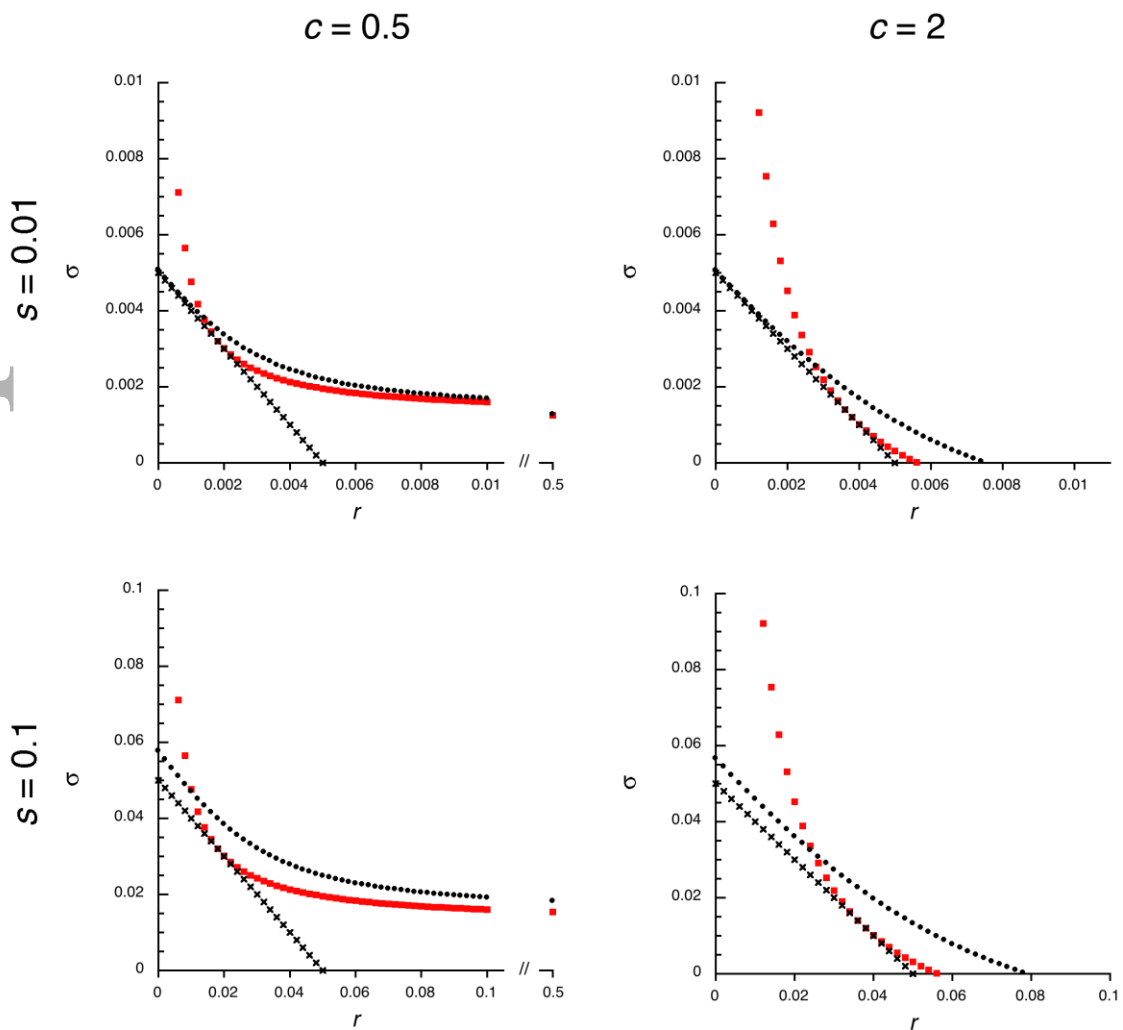


Figure 3. Results for the invasion of a population by male-benefit alleles in a situation when  $s = 0.01$ ,  $h = 0.5$  and  $c = 1$ , so that  $|1 - h(1 + c)| \gg s$ , the condition for the validity of equation (2), is not met.

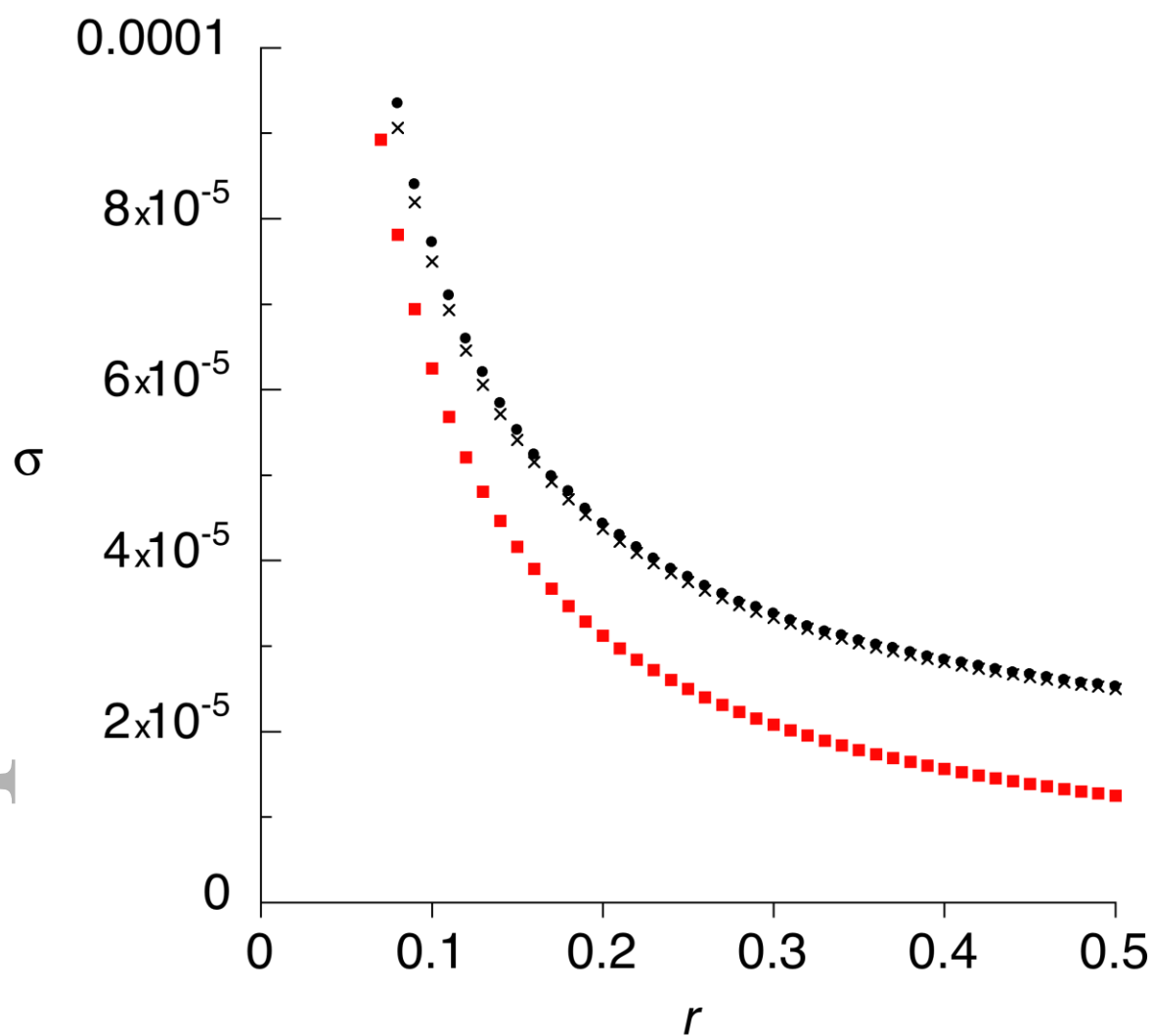
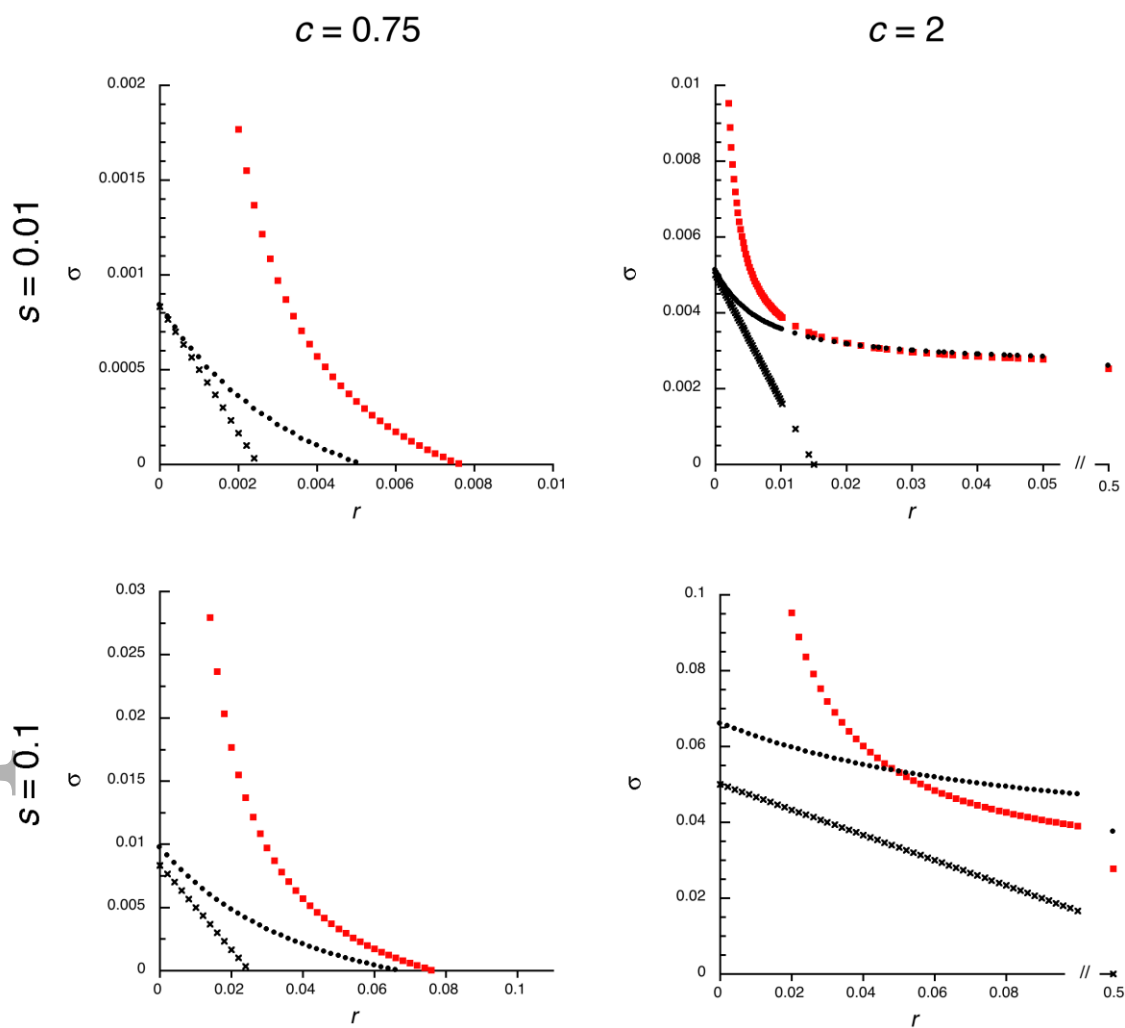




Figure 4. Results for the invasion of a population by female-benefit alleles. The figure is organised and labelled as in Figure 2, but the two values of the  $c$  parameter are 0.75 and 2, because with  $c = 0.5$ , the mutant never invades.



## Appendix

### ANALYSIS OF THE CHARACTERISTIC EQUATIONS FOR INVASION BY MALE- OR FEMALE BENEFIT MUTATIONS

#### Male-benefit case

Following Bull (1983, pp. 265-269), if second-order terms in the frequencies of genotypes carrying the invading allele  $A_2$  are ignored, we obtained a 3-dimensional matrix describing the recursion relations. The characteristic equation of this matrix has the form

$$\lambda^3 + p\lambda^2 + q = 0 \quad (\text{A1})$$

where

$$p = -\frac{(1-r)(1-hs)}{(1-s)} - \frac{(1-hsc)}{2} \quad (\text{A2a})$$

$$q = \frac{(1-hsc)(1-hs)^2(1-2r)}{2(1-s)^2} \quad (\text{A2b})$$

Writing  $\sigma = \lambda - 1$ , and including terms in  $\sigma^2$  but neglecting those in  $\sigma^3$ , equation (A1) can be reduced to a quadratic equation:

$$(3+p)\sigma^2 + (3+2p)\sigma + 1 + p + q = 0 \quad (\text{A3})$$

This has the following relevant solution, which is accurate to terms of order  $s^2$ .

$$\sigma = \frac{-(3+2p) + \sqrt{(3+2p)^2 - 4(3+p)(1+p+q)}}{2(3+p)} \quad (\text{A4})$$

This can be approximated further as follows, noting that the term under the square root sign can be written as  $(3+2p)^2 [1 - 4(3+p)(1+p+q)/(3+2p)^2]$ ; this is close to  $(3+2p)^2$ . Taking the first two terms in the Taylor series for the bracketed term, we obtain the second-order approximation

$$\sigma \approx -\frac{(1+p+q)}{(3+2p)} - \frac{(3+p)(1+p+q)^2}{(3+2p)^3} \quad (\text{A5})$$

This can be approximated to  $o(s^2)$  by using the second derivatives of  $p$  and  $q$  given in the ESM. Using the argument described there, the first term on the right of equation (A5) can be written as

$$\frac{s}{2}[1 - h(1 + c)] + \frac{s^2}{2}(1 - h)(1 - hc) + \frac{s^2}{4r}[1 - h(2 + c) + h^2(1 + c + c^2)] + o(s^2)$$

The second term can be approximated by

$$-\frac{s^2}{8r}(3 + p)[1 - h(1 + c)]^2 + o(s^2) = -\frac{s^2}{16r}(3 + 2r)[1 - h(1 + c)]^2 + o(s^2)$$

The final second-order approximation for  $\sigma$  is obtained by adding these two terms, yielding equation (1) of the main text.

The behaviour of  $\sigma$  as a function of  $r$  when  $r$  is small can be investigated by using implicit differentiation of equation (A3) to determine  $\partial S / \partial r$ . This gives

$$\frac{\partial S}{\partial r} = -\frac{(1 + 2S + S^2)\frac{\partial p}{\partial r} + \frac{\partial q}{\partial r}}{2(3 + p)S + 3 + 2p} \quad (\text{A6})$$

Using equations (A2), and the expressions for  $\sigma$ ,  $p$  and  $q$  when  $r = 0$ , the numerator of the fraction on the right hand side for  $r = 0$  can be written as

$$\frac{[2s(1 - h)(1 - s) + s^2(1 - h)^2 + (1 - s)^2 - (1 - s)(1 - hs)(1 - hsc)](1 - hs)}{(1 - s)^3}$$

For small  $s$ , the leading term in this expression is  $s[1 + h(c - 1)]$ . The leading term in the denominator is equal to  $3s(1 - h)$ . This yields equation (4) of the main text.

A similar procedure can be used to determine  $\partial S / \partial r$  when  $r = 1/2$ , which provides a check on the approximations used above. In this case, the selection coefficient for an invading  $A_2$  allele is  $1/2 s[1 - h(1 + c)] + 1/2 s^2(1 - h) + o(s^2)$ . Substituting this into equation (A6), together with the appropriate expressions for  $p$  and  $q$ , we obtain the final result:

$$\left(\frac{\partial S}{\partial r}\right)_{r=1/2} = -\frac{s^2}{4}[1 - h(1 - c)]^2 + o(s^2) \quad (\text{A7})$$

This agrees with the expression found by differentiating equation (1) at  $r = \frac{1}{2}$ .

### Female-benefit case

This case gives the following terms in the characteristic equation (A3):

$$p = -(1-r)(1-hs) - \frac{(1-hsc)}{2(1-sc)} \quad (\text{A8a})$$

$$q = \frac{(1-hsc)(1-hs)^2(1-2r)}{2(1-sc)} \quad (\text{A8b})$$

The derivatives of  $p$  and  $q$  used for the approximations for the selective advantage  $\sigma$  of  $A_2$  are given in the ESM. Using a similar argument to that for the male-beneficial case, the approximations shown in the main text can be derived.